

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

T S1/5/1

1/5/1

DIALOG(R) File 351:Derwent WPI

(c) 2004 Thomson Derwent. All rts. reserv.

011357387 \*\*Image available\*\*

WPI Acc No: 1997-335294/199731

XRPX Acc No: N97-278176

Coding of electronic text document to be transferred by computer system -  
embedding data into electronic document by modulating spaces between  
words, and modulating length of space preceding word in line of text  
relative to length of space following word with value of signature data  
bit embedded into text data

Patent Assignee: MATSUI K (MATS-I); OKI ELECTRIC IND CO LTD (OKID )

Inventor: HORINO N; MATSUI K; SUTO M

Number of Countries: 005 Number of Patents: 003

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
EP 782322	A2	19970702	EP 96309154	A	19961216	199731 B
JP 9186603	A	19970715	JP 95343295	A	19951228	199738
US 5861619	A	19990119	US 96767229	A	19961213	199911

Priority Applications (No Type Date): JP 95343295 A 19951228

Cited Patents: No-SR.Pub

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
EP 782322	A2	E	10	H04N-001/32	

Designated States (Regional): DE FR GB

JP 9186603 A 7 H03M-007/14

US 5861619 A G06K-007/10

Abstract (Basic): EP 782322 A

The method prepares an electronic document (1) carrying first text data (33) including words forming a line with a space in between them. Bits of a second text data (31) to be embedded in the text document are prepared. The length of the first space (pi) is modulated preceeding a word in a line of text in the first text data.

The space is modulated relative to the length of a second space (si) following the word with a value of a bit of the second text data. The second text data is embedded into the first text data. Prior to modulation taking place, every other one of the words except the top and lasts words, in a line of the first text data together with the first and second spaces with respect to the word are selected.

ADVANTAGE - Codes electronic text document which is then not easily discernible to unauthorised entity. Prevents illegal duplication.

Dwg.3/3

Title Terms: CODE; ELECTRONIC; TEXT; DOCUMENT; TRANSFER; COMPUTER; SYSTEM;  
EMBED; DATA; ELECTRONIC; DOCUMENT; MODULATE; SPACE; WORD; MODULATE;  
LENGTH; SPACE; PRECEDE; WORD; LINE; TEXT; RELATIVE; LENGTH; SPACE; FOLLOW  
; WORD; VALUE; SIGNATURE; DATA; BIT; EMBED; TEXT; DATA

Derwent Class: T01

International Patent Class (Main): G06K-007/10; H03M-007/14; H04N-001/32

International Patent Class (Additional): G09C-001/00; H04L-009/18;

H04L-009/32

File Segment: EPI

?

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-186603

(43) 公開日 平成9年(1997)7月15日

(51) Int.Cl. <sup>6</sup>	識別記号	序内整理番号	F I	技術表示箇所
H 0 3 M 7/14		9382-5K	H 0 3 M 7/14	A
G 0 9 C 1/00	6 1 0	7259-5J	G 0 9 C 1/00	6 1 0 D C3
	6 4 0	7259-5J		6 4 0 A C2
		7259-5J		6 4 0 D C2
H 0 4 L 9/18			H 0 4 L 9/00	6 5 1 C3

審査請求 未請求 請求項の数 3 O L (全 7 頁) 最終頁に続く

(21) 出願番号 特願平7-343295

(22) 出願日 平成7年(1995)12月28日

(71) 出願人 000000295

沖電気工業株式会社

東京都港区虎ノ門1丁目7番12号

(71) 出願人 000187312

松井 甲子雄

神奈川県横浜須賀町大津町5丁目57番地

(72) 発明者 堀野 直治

東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

(72) 発明者 須藤 正之

東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

(74) 代理人 弁理士 金倉 喬二

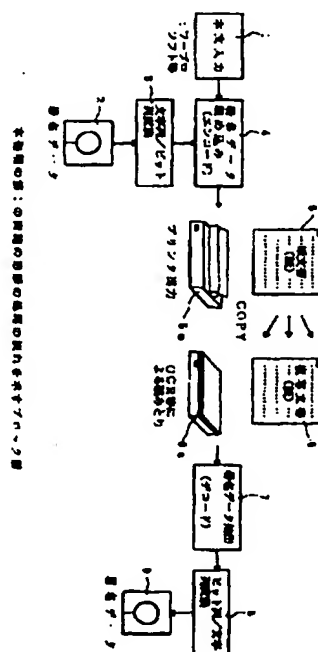
最終頁に続く

(54) 【発明の名称】 電子文書の単語間の空白部分の長さを利用した符号化および復号化方法、電子文書への署名情報の埋め込み方法、機密文書の暗号化方法

(57) 【要約】

【課題】 第三者に見破られることなく、多くの署名情報を電子文書内に埋め込むこと。

【解決手段】 署名データ埋め込み手段4は、ある単語の前にある空白の長さとその単語に引き続く空白の長さの組み合わせを1つの符号単位とし、前後の長さの比率を署名データのビット値に応じて変化させることで、前記符号に著作権を保護するための証明データを埋め込む。署名データ抽出手段7は、単語の前後の空白の長さの比率の違いに応じて署名データを抽出する。



## 【特許請求の範囲】

【請求項1】 コンピュータネットワークを介し電子形態で取引される文書内のある単語の前にある空白の長さ、その単語に引き続く空白の長さの組み合わせを1つの符号単位として符号化および復号化を行うことを特徴とする電子文書の単語間の空白部分の長さを利用した符号化および復号化方法。

【請求項2】 コンピュータネットワークを介し電子形態で取引される文書内のある単語の前にある空白の長さ、その単語に引き続く空白の長さの組み合わせを1つの符号単位とし、前後の長さの比率を署名データのビット値に応じて変化させることで、前記符号に著作権を保護するための署名データを埋め込み、前記署名データが埋め込まれた文書の複写文書から、該埋め込まれた署名データを、単語の前後の空白の長さの比率の違いに応じて抽出して認証を行うことを特徴とする電子文書への署名情報の埋め込み方法。

【請求項3】 コンピュータネットワークを介し電子形態で取引される文書内のある単語の前にある空白の長さ、その単語に引き続く空白の長さの組み合わせを1つの符号単位とし、この前後の長さの比率を機密文書のビット値に応じて変化させることで、前記符号に機密文書を埋め込んで暗号化することを特徴とする機密文書の暗号化方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、電子文書の符号化および復号化の方法に関するもので、特に、コンピュータネットワークを介し電子形態で取引される文書に署名情報を埋め込み、認証することにより、電子出版物の著作権を保護するための署名の埋め込み方法に関するものである。

## 【0002】

【従来の技術】近年、コンピュータネットワークの発展に伴い、電子的な方法で文書を迅速かつ経済的に配布することが可能になってきた。特に、電子ライブラリやデータベースの検索等により、多くの貴重な資料を不特定多数のユーザが利用する機会が増えるに従い、その著作権の保護が問題化している。活字を組み、本に印刷していた時代に比較して、電子文書は何の労力も要せず瞬時にして複写可能である。しかも、電子文書のコピーは原本と寸分たがわず、多くの人が無断で新本を入手したのと同じ結果となる。

【0003】このような電子出版物著作権の侵害を防止するために、電子文書に署名情報を挿入する方法として、以下の3つの方法が考えられている。第1の方法は、文書の行間隔を1行おきにわずかに上下にシフトし、その長短を情報に変換するものである。第2の方法は、テキストの中にある単語の位置を水平方向に前後に移動させることにより、原本との差異を情報に変換する

ものである。第3の方法は、特定の文字をわずかに変形させて情報を埋め込むものである。

## 【0004】

【発明が解決しようとする課題】しかしながら、上述した従来の第1の方法では、埋め込み情報量が1ページ内のテキストの行数に制限されること、また、電子編集時に等間隔でないことを見破られることがある等の問題がある。また、第2の方法では、常に原本との照合が不可欠であり、チェックに不便である。また、英文ワードプロセッサの多くは、1行単語数および文字数が、その行にうまく入らないと、次行に送ったり、あるいは字間を詰めたりする機能を有しており、このため、電子文書をハードコピーとして出力すると、しばしばこの機能により原文との不一致が発生してしまう問題がある。

## 【0005】

【課題を解決するための手段】上述した課題を解決するため、本発明は、コンピュータネットワークを介し電子形態で取引される文書内のある単語の前にある空白の長さ、その単語に引き続く空白の長さの組み合わせを1つの符号単位とし、前後の長さの比率を署名データのビット値に応じて変化させることで、前記符号に著作権を保護するための署名データを埋め込み、前記署名データが埋め込まれた文書の複写文書から、該埋め込まれた署名データを、単語の前後の空白の長さの比率の違いに応じて抽出して認証を行うことを特徴とする。

## 【0006】

【発明の実施の形態】図1は本発明の第1の実施の形態における署名情報の埋め込みおよび抽出処理の流れを示すブロック図である。図において、1は本文入力手段で、この本文入力手段1を用いて署名埋め込み対象となる重要文書の入力を行う。ここで、本文入力手段1は、特別な機能を持たない市販のワープロやPC上のワープロソフト等でよい。2は前記本文入力手段1で入力された重要文書に埋め込まれる署名データを格納した署名データ格納手段である。署名データには、契約書に示す記号情報を用いて筆者の氏名、契約番号、配布時間および使用条件等のデータが含まれる。3は文字列/ビット列変換手段で、署名データを文字列からビット列に変換する処理を行う。例えば、筆者の氏名等を2バイト漢字コードに変換し、さらに冗長コードを追加して、復号時に誤り検出および訂正を可能とする。

【0007】4は署名データ埋め込み手段で、前記本文入力手段1で入力された重要文書に、前記文字列/ビット列変換手段3でビット列に変換された署名データを所定の規則で埋め込む処理を行うものである。埋め込み方法については、図2を用いて後述する。5は署名データが埋め込まれた重要文書をプリンタ5aで打ち出した原文書である。ここで、プリンタ5aには特別な機能の追加はなく、市販のプリンタで良い。6は前記原文書5を複写機により不正に複写した複写文書で、この複写文書

6を、OCR6a等により、単語間空白情報として署名データを含んだ電子文書に復元する。なお、複写文書6を電子文書に復元するには、OCR6aを用いて文字認識しなくても、文書内の単語の位置情報が求められる装置があればよい。

【0008】7は署名データ抽出手段で、前記OCR6a等で復元した単語間空白情報として署名データが埋め込まれた電子文書から、該署名データを所定の規則で抽出する処理を行うものである。抽出方法については、図2を用いて後述する。8はビット列／文字列変換手段で、前記署名データ抽出手段7で抽出したビット列で表されている署名情報を、文字列に変換する処理を行う。9は前記ビット列／文字列変換手段8で文字列に変換された署名データを格納する署名データ格納手段である。

【0009】次に、署名データの埋め込みおよび抽出方法について説明する。図2は第1の実施の形態における署名データの埋め込み処理の詳細を示す説明図である。原文書5のテキスト行は、単語21、22・・・の並びから構成されている。ここで、本実施の形態は、英語のように単語の区切りがはっきりしている言語に適用するものである。そして、テキスト行のある1つの単語iに注目する。このとき、自身の前あるいは後ろの単語がないので、冒頭と末尾の単語は除く。図2では、単語22、24、i等、網掛けで示す位置の単語である。その単語の前にある空白長さを $p_i$ 、その単語に引き続く空白長さを $s_i$ とする。このとき、 $(p_i, s_i)$ の組を1つの符号単位と考え、署名データ埋め込み手段4では以下に示す規則を定義する。

【0010】＜埋め込み規則＞埋め込むべき署名データのビット値が、

(a) 0ならば、

$$p_i = (1 - \rho) \cdot (p_i + s_i) / 2$$

$$s_i = (1 + \rho) \cdot (p_i + s_i) / 2$$

(b) 1ならば、

$$p_i = (1 - \rho) \cdot (p_i + s_i) / 2$$

$$s_i = (1 + \rho) \cdot (p_i + s_i) / 2$$

とする。ここで、 $\rho$  ( $0 < \rho < 1$ )を偏移度と呼ぶ。また、長さの単位は、多数回の複写による画像の劣化およびOCR6a等による空白読み取り誤差を考慮し、量子化する。この手順を1語おきに繰り返すならば、署名データのビット列を容易にテキスト行の上に写像することが可能となる。

【0011】次に、単語間空白情報として署名データを含んだ電子文書からこの署名データを抽出するために、署名データ抽出手段7では以下に示す規則を定義する。

＜抽出規則＞対象となる単語iの前後空白長さが $p_i, s_i$ で、

(A)

$$p_i = (1 - \rho) \cdot (p_i - s_i) / 2$$

$$s_i = (1 + \rho) \cdot (p_i - s_i) / 2$$

ならば、抽出署名ビットの値は0

(B)

$$p_i = (1 - \rho) \cdot (p_i - s_i) / 2$$

$$s_i = (1 + \rho) \cdot (p_i - s_i) / 2$$

ならば、抽出署名ビットの値は1

とする。ここで、 $\rho$  (偏移度)は単語iによらず一定である。この手順を1語おきに繰り返すならば、署名データを示すビット列を容易にテキスト行の上から抽出することが可能である。

【0012】このように、本発明の第1の実施の形態では、署名データを抽出する際に、ある単語の前後の空白 $(p_i, s_i)$ の比率により署名データのビット値が0か1かを判断している。このため、 $(p_i, s_i)$ の相対的な関係により0か1かを判断でき、復号の際に原本と照合する必要がない。また、テキスト画面上に埋め込み可能な署名データのビット数はそのテキストの単語数の約半分となる。なぜならば、単語間の空白長さの組み合わせで情報を表現しているからである。したがって、A4版1ページに40行、各行20語程度のテキストでは、約400ビットの署名データを埋め込むことが可能であり、埋め込み情報量が多いものである。

【0013】次に、電子文書のある単語の前後の空白の組を1つの符号単位とすることを利用して、機密文書を暗号化する実施の形態について説明する。例えば、新聞記事のような任意の文書の単語の前後の空白の組を1つの符号単位とする。そして、ビット列に変換した機密文書の各ビット値に応じて、上述した単語の前後の長さの比率を変化させる手順を繰り返すことで、前記符号に機密文書を埋め込んで暗号化する。機密文書を復号する場合は、機密文書が埋め込まれた文書から、単語の前後の空白の長さの比率の違いに応じてビット列を抽出し、これを文字列に変換することで機密文書を得る。

【0014】

【発明の効果】以上説明したように、本発明は、電子文書中のある単語の前後の空白の組を1つの符号単位とし、前後の長さの比率を署名データのビット値に応じて変化させることで、前記符号に著作権を保護するための署名データを埋め込むこととしたので、署名符号化された文書であることを第3者が見破るには、かなり高い技術と努力を必要とし、困難である。また、全てのページ、全ての空白に繰り返し署名データを異なる組み合わせにして埋め込んでおくことにより、悪意を持つ不正コピー者に消去のための多大な労力と高価なコストを課すこととなり、不正コピーの防止効果を上げることができるとなる。

【図面の簡単な説明】

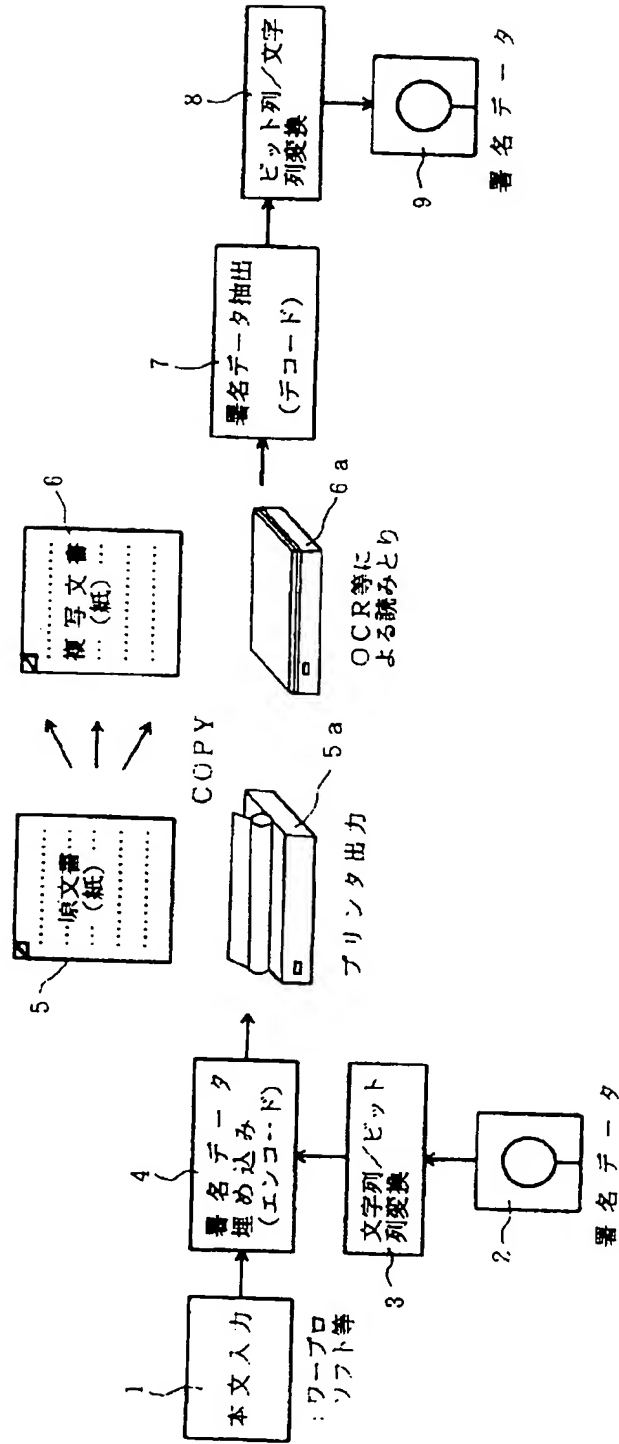
【図1】本発明の第1の実施の形態における署名情報の埋め込みおよび抽出処理の流れを示すブロック図

【図2】第1の実施の形態における署名データの埋め込み処理の詳細を示す説明図

## 【符号の説明】

- |                |                |
|----------------|----------------|
| 1 本文入力手段       | 4 署名データ埋め込み手段  |
| 2 署名データ格納手段    | 7 署名データ抽出手段    |
| 3 文字列・ビット列変換手段 | 8 ビット列・文字列変換手段 |
|                | 9 署名データ格納手段    |

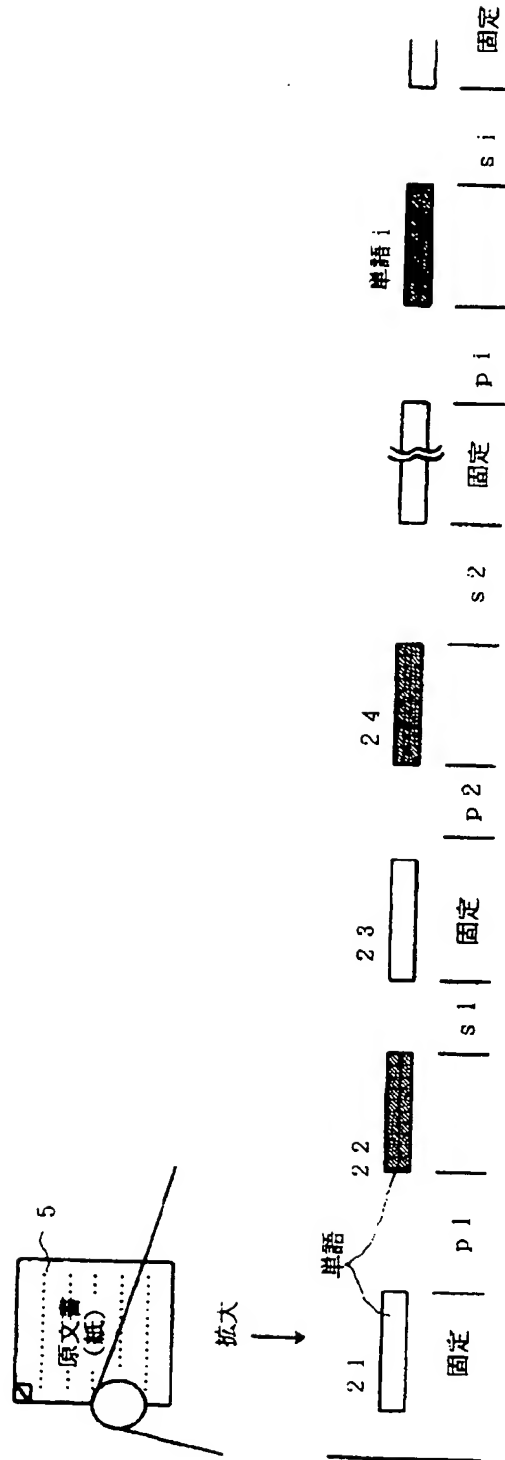
【図1】



本発明の第1の実施の形態の処理の流れを示すブロック図

(6)

【図2】



p i : 対象単語前部空白長さ

s i : 対象単語後部空白長さ

(p i, s i) : 符号化単位

第1の実施の形態における署名データ埋め込み処理を示す説明図



( 7 )

特開平 9 - 1 8 6 6 0 3

フロントページの続き

(51)Int.Cl. <sup>6</sup>	識別記号	片内整理番号	F I	技術表示箇所
H 0 4 L 9/32			H 0 4 L 9/00	6 7 5 A C2

(72)発明者 松井 甲子雄  
神奈川県横須賀市大津町 5 - 57